DOCUMENT RESUME

ED 168 167 EA 011 445

TITLE	Minimum Competency. Research Report 1978-3.
INSTITUTION	National School Boards Association, Washington, D.C.
PUB DATE	78
NOTE	39p.
AVAILABLE FROM	National School Boards Association, 1055 Thomas Jefferson Street, N.W., Washington, D.C. 20007 ($7.50; quantity discounts; shipping and handling charges will be added to billed orders)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Academic Achievement; Basic Skills; Board of Education Role; Elementary Secondary Education; *Federal Regulation; *Government Role; Government School Relationship; *Minimum Competency Testing; Policy Formation; Program Development; School District Autonomy; Testing Problems; Testing Programs

ABSTRACT

This four-part report deals with minimal competency testing with a focus on test development and the federal role in competency testing. The first section presents highlights of a 1978 conference on "Achievement Testing and Basic Skills" chaired by Secretary of Health, Education, and Welfare Joseph Califano. The second section contains the committee report "Improving Educational Achievement" issued by a nine-member committee headed by Stephen K. Bailey. This report, as well as the conferees at Secretary Califano's conference, advised the federal government against entering the business of developing national tests. Relying on the report, Califano recommended to the conference that the federal role be that of research, technical assistance, and as a clearing house for information. The third section of the publication presents the results of a survey of school board presidents about issues in competency testing. The results indicate that school boards would be vehemently opposed to the federal government's creation of national tests. The fourth section of the document is made up of recommendations concerning competency testing that the Education Commission of the States (ECS) has made to state and local governments. The ECS urges a joint effort by state and local authorities in the establishment of testing programs. (Author/JM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

NSBA NATIONAL SCHOOL BOARDS ASSOCIATION

Research Report 1978-3

MINIMUM COMPETENCY

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

NSBA

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ED168167

EA 011 445

Minimum Competency

National School Boards Association

Margaret Buvinger, President
Thomas A. Shannon, Executive Director

James Betchkal, Associate Executive Director, Office of Program Services
Lee VanBremen, Assistant Executive Director, Special Services Section
Crystal Kuykendall, Director, Constituency Coordination Department
James Mecklenburger, Director, Development Department
Mary Ellen Sera, Acting Director, Promotion Department

Research Report 1978-3
Minimum Competency
Price of Report: Single copy $7.50. Discounts on quantity orders: 10-49 copies, 10%; 50-99 copies, 15%; 100 or more copies, 20%. Shipping and handling charges will be added to billed orders.

Reproduction: No part of this Research Report may be reproduced in any form without prior written permission from NSBA. Address communications to the Direct Affiliate Program, National School Boards Association, 1055 Thomas Jefferson St., N.W., Washington, D.C. 20007.

Information reported in NSBA Research Reports does not necessarily reflect official viewpoints of the Association.

Copyright © 1978 by National School Boards Association.
All Rights Reserved.

Contents

Foreword v

Introduction 1

Highlights of the Commissioner's Conference, *Achievement Testing and the Basic Skills* 3

"Improving Educational Achievement" 7

Recommendations from the Education Commission of the States 19

The NSBA "Minimal Competency Testing" Survey 25

Foreword

Last winter, in *Back-to-Basics* (NSBA Research Report #1978-1), we observed that back-to-basics signals a shift in the education climate; that climate has since gotten much hotter. If we have heard slightly less about "back-to-basics," we are hearing far more about testing—especially about the use of tests to establish standards for grade-to-grade promotion or for high school graduation.

The pressure for minimum competency tests has often surfaced in Washington these last few months as a desire for the federal government to create a national test. School boards are opposed to this, as are most educators. The Carter administration—in the persons of Secretary of Health, Education, and Welfare Joseph Califano, Assistant Secretary for Education Mary Berry, Commissioner of Education Ernest L. Boyer, and Director of the National Institute of Education Patricia Graham—has opposed development of federal tests. By sponsoring the HEW "Achievement Testing and the Basic Skills" conference and the National Academy of Education "Improving Educational Achievement" report, which are reviewed extensively in this report, they have focused the intellect and the muscle of the education community against any such federal tests.

The issues are greater than who should design such tests. Some of the other questions surrounding competency testing are reviewed in this report. Perhaps the most influential expression of these concerns is in the Policy Recommendations of the Education Commission of the States (ECS). After a year of studying the flurry of national activity, ECS makes very cautious recommendations, calling on state and local governments to go slowly, to move deliberately, and to understand the limitations of the tests they wish to establish.

We learned from our own survey of school board members, prepared for our delegation to the "Achievement Testing and Basic Skills" conference, that school boards have some sympathy for the basic premises behind competency testing but great concern about who designs such tests and about how such tests are to be used. By and large, board members would want neither to penalize the schools nor to stigmatize students by such testing, but rather would want any such testing program to improve teaching and learning.

Thomas A. Shannon
Executive Director
National School Boards Association

Introduction

Competence. Competency. Minimum. Minimal. Tests. Testing. Standards.

These are the words which, in 1978, are forming local, state, and national policy about schooling. Legislators in almost every state have considered or have established some kind of statewide testing program. As we reported in *Back-to-Basics* (NSBA Research Report #1978-1), "the back-to-basics movement is epitomized by minimal competency testing." But there are more motives afoot than a swing "back" to traditional practices and concepts. The motives of one or another party seeking to establish competency testing include:

- To assure that only competent students graduate.
- To diagnose learning failures and to improve schooling.
- To make schools more accountable for results.
- To prove that schools are doing well; to silence angry critics.
- To prove that schools are doing poorly; to support angry criticism.
- To "motivate" teachers, students, boards, or administrators to perform better.
- To justify new funds for remedial instruction.
- To restrain the costs of education by eliminating "frills."
- To create quantifiable public information to aid policy makers, including governors, state departments, state boards, and legislatures (and, perhaps, to justify state rather than local intervention).
- To get back to the business of competing with the Russians in educational quality.
- To sell tests.

In March, Secretary of Health, Education, and Welfare Joseph Califano chaired a national conference on "Achievement Testing and the Basic Skills" at which the nation's education leaders were to advise the federal government on its role in the area of competency testing.

The conference (March 1-3) also marked the release of a report of a blue ribbon committee appointed by Assistant Secretary for Education Mary Berry. A panel of nine leading education scholars, chaired by Stephen K. Bailey, under the sponsorship of the National Academy of Education, responded to eight important questions about competency testing and basic skills.

This report, "Improving Educational Achievement," as well as the conferees at Secretary Califano's conference, advised the federal government that it should not be in the business of developing national tests. Relying on the report, Califano told the conference that the federal role will be one of research and technical assistance and as a clearinghouse for information.

"Improving Educational Achievement" has been both influential, because Secretary Califano has relied heavily on it, and controversial in its recommendations. It is reprinted in this report (pp. 7--17).

NSBA was invited to bring a delegation of eight persons to the Secretary's conference. In order to ensure that this delegation would have the widest possible view of school board concerns, NSBA Research conducted a survey among school board presidents about issues in competency testing. A

report of this survey is on pp. [illegible] 35. It demonstrates that school boards would be vehemently opposed to the federal government creating national tests. Board members would prefer that such tests be constructed locally.

Also, during 1977-78, the Education Commission of the States has devoted a year of policy study to the competency testing movement. It has produced a series of conferences, several important publications, and, most important, a set of policy recommendations to state and local governments.* These recommendations, approved June 17, are bound to be influential and read widely because ECS counts as its audience the state-level leadership in education in 49 states (governors, legislators, state boards and state departments, and the education organizations).

ECS says that competency testing must be approached with considerable caution. It urges a joint effort by state and local authorities in the establishment of testing programs. It also urges attention to the design of tests, so that such tests are not charged with trying to accomplish things they cannot. In addition, ECS urges that tests not be the only criterion applied for high school graduation. ECS's recommendations are reproduced on pp. 19–24.

* Highlights of ECS's work are contained in the May issue of *Phi Delta Kappan*. Complete reports are available from the Education Commission of the States, 300 Lincoln Tower, 1860 Lincoln Street, Denver, Colorado 80295.

Highlights of the Commissioner's Conference, *Achievement Testing and the Basic Skills*

How poorly are the schools performing? Is the minimum competency movement to be taken at face value, or is it merely a naive reaction to the public's frustrations about education? Is there a federal role in these matters? Can testing accomplish what the advocates of competency testing want it to accomplish?

In "Tests and Schooling," the keynote address to the conference, Harold Howe II, formerly U.S. commissioner of education and currently vice president for education and research, the Ford Foundation, posed these questions in provocative and surprising ways:

> If anyone had told me ten years ago that such a meeting as this would come about, I would have bet ten to one against it because of the belief that testing in the schools was not the business of the federal government. But here we all are—educators, parents, testing experts, civil rights advocates, local, state, and federal officials, and others.
>
> The reason most of us are here is not really our abiding interest in the theory and practice of achievement testing. We are here because many Americans are deeply concerned about their schools. They think that the schools are doing an inadequate job of helping children to learn, especially in fundamental skills that are the tools of more advanced learning. Whether they are justified in thinking that, no one can say with certainty. But there can be no question about the pervasiveness and strength of the feeling that American schools suffer shortcomings.
>
> The origins of these feelings are complex, and people who agree that schooling needs improvement don't agree at all on the roots of the problems, on what needs to be done about them, or about how serious they are. Using a medical analogy, some think that the condition is like a cold, which will go away if simple remedies are taken; others would call it flu, requiring more serious treatment; and some would put the patient in the hospital for oxygen and strong drugs because they think he has double pneumonia. Because no one really knows what the condition is and because our capacity to measure it is limited, there is some tendency for all of us to behave as if it might be pneumonia. Only then, it is argued, can we be sure that we have done everything possible.
>
> I think that compared with the rest of the world, schools in the United States are doing a good job. Since late 1968, when President-elect Nixon suggested I vacate the United States Commissionership of Education, I have had considerable opportunity to observe education outside the United States, as well as at home. . . . There is much we can learn from schools elsewhere and much they can learn from us; there is plenty of room for improved practice in American schools, and some current evidence that particular aspects of schooling require special attention.
>
> But there is no reason to think that American education is in such a perilous state as some would have us believe. In other words, I don't think we have pneumonia, and I believe that treating the patient as if he did could do him unforeseen damage. While test scores have declined on many tests since 1960, some tests report unchanged scores, and others show gains.
>
> Secondly, I call attention to the very great achievement of American schools in bringing into their classrooms almost all the children of almost all the people. . . . The story of American education in the twentieth century is one of great achievement. We have built a diverse galaxy of public and private schools, and we have rid ourselves of a legally supported and racially defined dual system of schools. In 1900 about 6 percent of youth finished high school; today the comparable figure is 80 percent and about half of them go on to further education. . . .
>
> We have forged a broad public commitment to the importance of education in the life of the re-

public. Despite some fraying at the edges, that commitment has not really wavered. For evidence look to the current public concern about levels of achievement among students. Americans want their schools to do more than just enroll large numbers of students. They want the students to learn something of value to them and to society.

There always will be argument about exactly what should be learned. But there is a powerful consensus that it should include improved capacity to know and use both our common language, English, and the universal language, mathematics. Modern achievement testing is sending out signals that the schools are falling short of these goals. To keep things in balance, these signals should not be allowed to dim the achievements of the schools in these important skill subjects. If one were able to select at random half a million 18-year-olds from 1938 and another half a million in 1978—high school graduates and non-graduates—and give them almost any achievement test, the 1978 group would perform significantly better.*

. . . We should value the great diversity of institutions we have encouraged, partly by happenstance and partly by design. This diversity is especially clear in higher education with its great variety of public and private institutions. In schools the scene is more dominated by the public sector, and the parochial and independent schools are numerically less significant. But they are important just the same. Within the public schools there is diversity with respect to the nature of schools and probably more variety than there ought to be in quality among those with similar missions. This diversity starts with the prerogatives of local school boards and is augmented by the existence of fifty separate state systems. It is powerfully reinforced by the nature of federalism in the United States and by historical traditions.

Even within the same school system and under one board of education, individual schools and teachers within those schools have important freedom in varying instruction and materials to fit their pupils. Our basic resource for improving quality in the public schools lies in helping teachers and principals to do a better job, and most of them would welcome such assistance. Changes in tests, textbooks, or promotion requirements are secondary lines of action that will not make much difference without the interest and support of teachers.

* Results of a study by Indiana University and the Indiana State Department of Education, announced April 15, showed that sixth graders in 1976, taking the same test that 6th graders took in 1943–44, "outscored their earlier counterparts on every subtest and on total median score."

. . . I find in present legislative proposals before Congress for national achievement testing an intent that opens the door to a nationalized school system. This concept is anathema to the traditions of American education. Federal initiatives can make some contributions to improving learning of fundamental skills, but it does not make sense to take precipitous action on nationally sponsored achievement testing. Not only would such action alter the nature of our country's diverse education arrangements, it would also be of doubtful use in improving what schools do to build basic skills in children.

. . . There must be opportunities for teachers to learn more about using tests effectively, particularly as new tests are developed. Congress can act on this front by encouraging the in-service training of teachers. It could re-enact the Education Professions Development Act that was unfortunately allowed to lapse in the mid-70s. If we had it now, it would be an important tool for teacher training in the use of tests to improve basic skills. If any interest develops in reviving this legislation, it should be specially focused upon in-service training and should include the training of principals and other administrators as well as school board members.

No amount of giving tests to children will ever teach children to read and write better. . . . But if a change is to take place between a particular teacher and a particular child some change will have to occur in the teacher's repertoire of skills.

I have some concerns about the minimum competency bandwagon that is riding down the main street of American education today. Suppose we are successful in improving our national achievement test scores by warning kids, "You can't enter the ninth grade until you can read at the sixth grade level," or "You can't graduate from high school without eighth grade reading skills." No doubt such pressures will result in improved skills among some youngsters, but what do we do with the children who just don't make it, either because of the shortcomings of their schools or their own problems? Is American society ready to live with a growing number of young persons who are rejected by the schools between age fourteen and seventeen without credentials that will allow them a chance at a job and a decent life?

. . . This meeting should also consider whether tests are in some way unfair to minorities who score lower than the general population. . . . While on the subject of fairness to minorities, I have to raise the difficult question of whether the national need for improving basic skills performance has hidden within it overtones of racism. . . . Basic competency testing will certainly result in more educational delays and denials for minorities than for the rest of us. The only resolution of this difficulty I can accept is a clear commitment by

those who advocate basic competency testing to provide at the same time all the special additional services that may be needed to help less fortunate children meet the test requirements.

Before we enshrine the use of testing in legislative and administrative actions we must ask whether we really know enough about the measurement of cognitive ability. When legislators and high-level policy persons start making firm and detailed decisions in highly technical areas they can end up making mistakes. In no sense am I questioning the motives of those who would stipulate testing activities in national or state legislation. Clearly they want to help children by improving schools. Yet somehow, as I watch state legislatures legalizing Laetrile as a cancer cure, I can't help worrying about their blessings for certain forms of mental testing.

From what little I know about testing, I am inclined to think that the basic competency movement is rushing schools into wider use of tests, some of which are simply inappropriate for the broad purpose of improving basic competency. It takes several years to make a good test, try it out, and find what validity it has. That kind of effort is not reflected in competency tests that are hastily concocted in response to political directives.

. . . We will not make progress by sloganeering, whether we shout "back to basics" or "forward to fundamentals." We will not make progress by creating scapegoats. . . . The controversy brought on by the back-to-basics movement and the accompanying rapid move to basic competency testing and to wide use of testing for assessment of schools has unfortunately polarized educators and the friends of education. But there is really no essential conflict between improving teaching and learning in fundamental skills and the interest of educators both in developing good citizenship and sound values and in offering to young people opportunities in such areas as art and music. There may be some extremists in the back-to-basics camp who want to use it for the ulterior purpose of cutting education costs by ridding the schools of important activities, which they regard as frills. But this is not characteristic of most of the people harboring a concern for improving learning skills. So I return to where I started—we are fortunate in the schools we have and fortunate also that we are dissatisfied with them.

Harold Howe remarked on the matter of local prerogatives and diversity within American education. In one of the few emotional moments of the conference, Ann Kahn, representing the national PTA, but herself also a member of the Fairfax County, Virginia, Board of Education, engaged in discussion with John Porter, Michigan's State Superintendent of Public Instruction. Ann Kahn, recognizing the close relationship of testing to curriculum, made the case that "the authority to design tests ultimately is the authority to alter curriculum, and that authority is best exercised at the local level. Neither the state," she said, "nor the federal government have any constructive basis for tampering with that process." Not so, said Porter. It is the state which has the responsibility, pays the costs, and harbors the expertise to operate a school system for all children and to test to see whether that responsibility has been met.

It is clear that one of the threads of concern that runs through the competency testing movement is whether local officials are believed to have the responsibility to test what is taught.

In a panel designed to exhibit the thinking of national political figures on the issue of competency testing, Representative Michael Harrington (D-Mass.), Senator S. I. Hayakawa (R-Cal.), and Senator Claiborne Pell (D-R.I.), Senate Education Subcommittee Chairman, each spoke about the heavy mail they receive from constituents who regard federal activity as the necessary solution to their local problems with schools. Senator Pell inquired of the audience whether anyone thought the idea of federally developed tests was wise. He found only four raised hands. Responding to this conspicuous lack of support, Pell said that a national test is a good idea whose time will yet arrive.

Secretary Califano, in his formal remarks, made clear that President Carter himself "has a deep personal interest" in the subject of testing. He read a greeting from the President which said he was "delighted" by the conference because "there is no greater challenge than the one facing our educational system—to ensure that all children learn, at the very least, to read, to write, and to compute. . . . I believe that achievement testing can and should play an important and constructive role in the educational system."

Califano then made these commitments:

- To establish within the National Institute of Education a new Office on Testing and Assessment. "Its principal role would be to help people with questions find places with answers throughout the nation."
- To conduct a series of regional conferences on testing, sponsored by the regional HEW offices.

- More research on testing, including a new priority in this area at NIE.

He dwelt on the major findings of the National Academy of Education report, "Improving Educational Achievement," stressing that evaluating its findings is a "fundamental responsibility" of state and local governments. "Local control of education is, and remains, an important national tradition. But the federal government can help."

During the conference, small discussion groups met to render advice to the federal government, and these discussions were summarized at the final meeting of the conference. According to Michael Kaplan, who rendered the summary, there were four overarching themes:

- There was widespread suspicion about the real purpose of the conference, heightened by Secretary Califano's making announcements of new federal initiatives.
- Conferees agree wholeheartedly that there should be no federal test, created directly or indirectly by the federal government.
- We must not lose sight of the big picture: Testing should be understood as but one element of evaluation; evaluation as one component of schooling; schooling as one component of learning; and learning but one component of improvement of society.
- Schooling is everyone's business. Discussions about testing or any education issue should include the active participation of the many constituencies represented at the conference.

There were three additional points of nearly universal agreement:

- Schools are not disasters. They need fine tuning and they need leadership, but they do not need overhaul.
- Access to education by all Americans is an important goal, part of an unfinished national agenda. Testing must aid in serving this goal, not make it more difficult.
- Whatever changes are made, they should be in the context of the good characteristics of American education we wish to preserve, things variously referred to during the conference as "autonomy," "diversity," "pluralism," and "local control."

About testing, it was widely agreed among conferees that the minimum competency movement is a manifestation of an age of consumerism and accountability in the society at large; that tests are social dynamite and can do irreparable harm as well as good; that the issue should not be more testing or less testing but better testing and better test usage; and, most important, educational goals should set the agenda for testing, not the other way around.

"Improving Educational Achievement"

This report represents an effort on the part of nine leading educators and scholars to answer eight policy questions about minimum competency testing and related concerns.

Consensus was sought; where members of the panel disagreed, these areas of disagreement were deleted from the final report. The report has subsequently been criticized on this procedure: that it represents no more than opinion, that nine different persons might have reached very different conclusions, that research might not support the opinions expressed, and that, in general, the report is not as authoritative as Secretary Califano made it appear.

Nevertheless, the strength of this document is its ability, in straightforward English, to state the issues and to express at least one important point of view in response to these issues. Secretary Califano, as he introduced the report, said, "We all may differ with some of its conclusions, but its major findings seem to me thoughtful and important."

In brief, "Improving Educational Achievement" finds that: schools bear considerable responsibility for declining scores on the Scholastic Aptitude Test; elementary schools need to focus more intensively on basic skills; more time needs to be given to basic skills instruction, and diagnosis of learning problems must be earlier and better; raising test scores per se *is a trivial goal, but increasing achievement is not; using tests for high school graduation requirements is "unworkable," "exceeds the present measurement arts," and will create more problems than it will solve; the school district or the school is the appropriate unit of accountability; there should not be national tests, mandatory or voluntary; and the apparent distinction between norm-referenced and criterion-referenced testing is more complex and murky than is generally stated.*

In November 1977 Secretary of Health, Education, and Welfare Joseph Califano, and Assistant Secretary for Education Mary Berry, reflecting widespread official as well as public concern about declining SAT scores and an alleged decline in the basic-skill competencies of American young people, asked the National Academy of Education to establish a panel on testing and basic skills. The Panel was composed of Stephen K. Bailey, Graduate School of Education, Harvard University (chairman); John B. Carroll, University of North Carolina, Chapel Hill; Jeanne Chall, Graduate School of Education, Harvard University; Robert Glaser, University of Pittsburgh; John I. Goodlad, University of California, Los Angeles; Diane Ravitch, Teachers College, Columbia University; Lauren Resnick, University of Pittsburgh; Ralph W. Tyler, Science Research Associates, Chicago; and Robert L. Thorndike, Teachers College, Columbia University.

The Academy Panel met with a group of staff experts from HEW on November 16, 1977. Assistant Secretary Berry attended that meeting. Subsequently, Assistant Secretary Berry, on behalf of the Task Force on Basic Skills and Quality Education, submitted to the National Academy of Education Panel a list of eight policy-related questions and asked that they respond within a two-month period. Each member of the Panel reacted

in writing to each question. The responses were then collated by Dr. Arthur E. Wise, formerly of NIE. His preliminary draft was circulated to members of the Academy Panel and to select HEW staff. A colloquium for the draft was slated for February 7, 1978, but had to be postponed because of weather. A smaller colloquium took place on February 16, 1978. The preliminary draft was then re-worked into the following report. It represents a distillation of the best thinking of the NAEd Panel. As is inevitable in a complex and evolving field, there emerged differences of opinion in the Panel itself. The body of the report represents areas of agreement.

The Panel is indebted to Assistant Secretary Berry and Deputy Assistant Secretary Joel S. Berke and his staff for a variety of facilitative and documentary services.

Stephen K. Bailey
Chairman

1. The findings of the Wirtz report* (On Further Examination) *are that many factors have been involved in the recent decline in test scores. Have there been, in your judgment, changes in schooling in the past two or three decades that have also contributed to the decline? Such factors as declining standards, too little instructional time, and changing curriculum have been mentioned. Is there evidence to bolster or to question these and related assertions?

For 14 years there has been a drop in the average score on the Scholastic Aptitude Test taken by high school students seeking admission to college. There has also been a consistent decline in recent years in the ability of 13-year-olds and 17-year-olds to write acceptable English.

These declines have understandably triggered both concern and analysis. American education has set itself a most difficult but necessary task: mass literacy of a high order. In analyzing our successes and failures, we must concern ourselves both with monitoring excellence and assuring those levels of minimum academic achievement that all students should attain. Our own work suggests that ever greater numbers of people have reached a higher level of literacy than ever before in our history. At the same time, test score declines suggest a flattening out of the educational pyramid. Thus we see three problems which are different but not unrelated: first, those who are left behind by rising standards of minimum literacy; second, declining achievement among the college bound; and third, the overall achievement of American schools. The evidence suggests cause for concern about academic performance of each of these groups.

Whatever non-school factors may have been involved (e.g., TV, parental permissiveness, the size of the age cohort in high school), there is growing evidence that shifts in policy, expectations, and behavior within schools themselves have contributed substantially to the documented decline in writing skills and in aptitude test scores. The Wirtz committee, commissioned by the College Entrance Examination Board to study the test score decline, identified in 1977 some salient within-school factors. The NAEd Panel believes that four within-school factors have been particularly important in causing declines in writing skills and SAT scores.

First, the proliferation of courses. For whatever reasons, many of them originating as external pressures upon the schools, the past 20 years have seen a proliferation of subject-matter courses available to pupils in the upper levels of elementary school and in junior and senior high schools. With these increased options, students have been able to choose courses less demanding than the traditional regimen of English, math, science, and social studies (especially history). The proliferation of non-intellectually rigorous courses has, unfortunately, been accompanied by a decline in the intellectual standards demanded by formerly rigorous, still required courses.

Second, confusion about the appropriate role of teachers. The 1950s and 1960s were periods of mass questioning about the ends and means of education. Teachers were given new and often contradictory models of appropriate pedagogic behavior. "The discovery method," "the open classroom," "indi-

vidualized instruction," and "team teaching" became "in" words, but they were rarely accompanied by teacher training and professional development programs adequate for effective implementation. In effect, teachers were confused. They floundered between what they knew how to do and what they were more and more being expected to do but not prepared to take on effectively. There was very little good traditional teaching and a paucity of creative approaches to the new.

Third, slackening of "on-task" attention. Harnischfeger and Wiley have learned that time spent in school is directly related to school achievement. But even more important is the percentage of time actually spent by students applying themselves to scholastic tasks.

Some studies suggest that even in relatively "good" classrooms students are "on task" for only about 30 percent of the instructional day (and the instructional day in a "good" classroom is devoted to learning only 60 or 70 percent of the time).

Goodlad, Klein and Associates suggest that an extraordinary amount of school time is spent in transitions from class to class, routines, announcements, interruptions, and general control and discipline.

Fourth, a dismantling of opportunities for intensive study in selective academic environments at the secondary level. In city after city programs for the academically talented have been closed or forced to accept changed admission standards. Islands of excellence within our public schools are rarer today than 20 years ago. We believe that this has contributed to the decline in numbers of students performing at the highest levels on SAT and similar tests. We need to improve instruction in the elementary schools to the point where access to certain programs at the secondary school level can be competitive without resulting in economically and ethnically segregated classrooms.

Although comparative studies are hard to come by, the NAEd Panel is convinced that "on task" attention has slipped in the past two decades, and that more effective use of school time in whatever the curriculum calls for would be a most significant reform in American schools.

2. Official as well as popular statements are increasingly being made in defense of a return to "basic skills." Do you believe these statements are justified? Are "reading, writing, and numbers" an adequate definition of "basic skills"? If not, what are others? What curricular balance should be sought among the basic skills identified? What are the trade-offs in focusing energy and resources on improving student achievement in the basic skills areas?

Two realities underscore the importance of reemphasizing "basic skills," assuming this phrase includes a learning core in reading, writing, and numbers. First is the deplorable and declining quality of written composition among all sectors of young people in this country. Second is the heavy penalty that functional illiteracy exacts from the 20 percent of American youth who come from homes where parents have had little or no education—homes largely concentrated in inner cities and isolated rural areas.

The term "basic skills" at its periphery is subject to a variety of meanings. Most people would agree, however, that reading, writing, and numbers are at the cognitive core. Experts disagree about how developed these skills must be in order to have individuals claim that they are truly competent in them. And many people would argue that certain affective (i.e., socially oriented) skills are essential if children are to become useful citizens and are to be prepared adequately for their various roles in society. The Academy Panel would only underscore the proposition that the distinctive role of the school lies in the development of cognitive functioning. It is the only social institution that has the development of cognition as its primary purpose. The basic skills are the building blocks for the more complex skills of scientific and value analysis; technological and occupational coping; historical, psychological, and philosophical understanding; and political participation. Without the acquisition of the fundamental learning tools of reading, writing, and numbers, no education of value can proceed.* Furthermore,

* It should be noted that reading and writing implies facility with and growth in the related language skills of speaking and listening.

a complex technological and bureaucratic society demands a high level of literacy from all those who would participate usefully and successfully in its functioning. The standard is higher than it ever has been before, and should be, and more people are expected to reach it, and should be. In such circumstances, a renewed emphasis on basic skills at these higher standards seems to this Panel eminently justified.

More specifically, American education should be paying much more attention to doing a thorough job in the fundamentals of reading, writing, and numbers beginning with ages five to eight, but then expanding the curriculum so that these basic skills are utilized in a broader, more demanding curricular diet. It is at the higher grades that the question of curricular balance becomes important. It is not reasonable to think that secondary schools, for example, will challenge and retain young people on a diet of only reading, writing, and arithmetic—even though enriched curricula at higher levels must always provide for the continuous upgrading of basic-skill competencies. And where teachers in the lower grades have not fulfilled their educational responsibilities in the basic skills, teachers at upper-grade levels and in secondary schools must make up for these deficiencies even at the cost of a leaner curricular diet for the learner who requires additional instruction. Otherwise, students will suffer from progressive incapacity to cope in a highly competitive society.

This is not to say that basic skills are the be-all and end-all of education. Clearly they are not. They are essential but not sufficient. Schools bear a large share of the burden of developing advanced analytic, critical, and artistic skills. Oral skills as well as the basic language skills of reading and writing need cultivation. Schools share with other community institutions a responsibility for developing in young people various social skills. An effective curricular balance to accommodate these various public expectations must be carefully structured at the upper levels of schooling. But such a balance is unthinkable without adequate basic skills acquired in the primary and elementary years.

The issue of "trade-offs" in focusing energy and resources on improving student achievement is specious. If a good foundation has been set in basic skills, the further honing of such skills and the acquiring of an expanded universe of knowledge at higher grade levels are mutually reinforcing—at least for the well-taught student.

3. What are promising directions for schools to take to improve educational attainment levels?

The question seems to assume relatively simple universal answers. Actually there are no panaceas, even if there are promising findings and developments.

Extrapolating from research findings, the answer to the question of how schools can improve educational attainment lies in spending more time on those attainments we value.

There is a striking convergence of evidence that points to the role of time-on-task—"engaged time"—in improving performance in school subject matters. This line of evidence, however, needs interpretation, for it is subject to a great deal of oversimplification and to potentially regressive actions on the part of schools eager for quick (and cheap) increases in test scores. At first glance, the evidence seems to call for a return to group, teacher-led instruction, since this is a known way to keep most children at work a good percentage of the time. But the evidence suggesting that more time spent leads to stronger performance is, virtually without exception, evidence based on group means. It shows that, on the average, groups with more exposure will do better on tests than groups with less exposure. It does not say that every child does better simply by lengthening time on task. Two groups might not profit from this: the very bright, for whom the extra time is redundant (they might profit more from being allowed to move ahead in a program, or at least being left free to read and explore on their own) and the relatively slow, who may not follow what is going on in group instruction aimed at the "middle" and, thus, although present, not be really "on task" in an intellectual sense. These children need to work, too, but on something within grasp. Thus, some degree of individualization seems to be needed.

Learning problems need to be diagnosed in individuals at an early age and related remediation

needs to be administered before children fall hopelessly behind. In short, more attention needs to be paid to individual progress, not only to a group mean. In statistical terms, more attention must be paid to the fine points of distributions of scores and not only to means and medians. In the rush to show test score gains, the temptation to attend only to means, and to use the easiest and least expensive methods of raising average scores, may lead to giving up much of what has been painfully learned in the last 15 years in the way of individual planning and personalized teaching.

There is another potential loss in too quick a return to the old-fashioned group-organized instruction. That is a loss of a new and more open social environment, which has permeated many schools as more individual activity has become normal. Perhaps the "open classroom" movement pushed the matter too far for most people, but it is probably also the case that most Americans would regret having to return to the old "line them up in rows" authoritarian classroom in return for increasing test scores. Perhaps the trade-off is not required. If "time-on-task" evidence is analyzed coolly, it argues that any organization of schools that keeps time-on-task high will solve the achievement problem. If time-on-task can be increased while simultaneously giving children reasonable degrees of control over their own time and learning, both increased achievement and responsive social environments may be possible. Some computer-aided and other programmed instruction have provided at least partial positive clues to resolving this dilemma.

Other promising directions include the following:

—With regard to level of challenge, it appears that easy programs do not lead to as high achievement as more difficult ones. What the optimal level is is not known but important to know. Research is needed here.
—For beginning reading, a "code" emphasis seems to produce better results than a "meaning" emphasis.
—Early diagnosis and remediation are more effective than later intervention.
—Reading to children, particularly on a level above their own linguistic development, is related to further linguistic development and to reading development.
—A competent and energetic school administration is essential for pupil achievement—in setting standards of expectation, in providing reading specialists, in making proven instructional materials available, etc.
—In some subjects, the "mastery learning" formats of Benjamin Bloom and others show great promise—although their cost in adequate teacher training and the difficulties of enforcing extra-classroom remedial exercises should be squarely faced.
—In some settings, peer instruction has had a profound effect on educational achievement.

There are, however, no easy substitutes for well-trained and committed teachers, principals who are helpful to and supportive of staff, a disciplined atmosphere in schools, improved pupil nutrition and health, and ultimately the raising of the standards of living of the poor in our midst.

Are there differences that are related to income and minority status?

There is no evidence to suggest that anything fundamentally different is needed for low-income or minority children than for any other children. They, too (perhaps they especially), need to get down to work on academic tasks. Such success stories as we have with minority and low-SES groups all point to the power of a businesslike atmosphere in which lots of structured academic work gets done. The problems (overlooking individuals in the search for average gains) and the trade-offs will probably be much the same for these groups as for any others. It is indeed possible that as we become more refined in our efforts to teach more to everyone, different approaches will turn out to work better for different subgroups in the population. But we are doing so little real teaching now, especially to low-SES and minority children, that we can probably show enormous initial improvement by simply attending to the elementary task of doing more teaching and keeping expectations high. The fine tuning can come later. There is no surer way to keep minority children from a fair chance than teaching them something different from what everyone else is being taught (or not teaching them much of anything at all). Note that this does not imply that we should pay no attention to cultural differences of children. If children speak no English, there is no point trying to teach them to read English as if they did. Their special situation must be realistically dealt with—perhaps by longer hours and more weeks in school each year since there is so much more that they have to learn.

What role should standardized tests play? How should they interact with curriculum and school programs?

Tests can play several different roles. One is as a means of public accountability, a way in which everyone interested can keep track of whether the schools are actually doing much teaching. The other is in helping to manage the instructional process, by either grouping students or diagnostically analyzing their competence so that instruction can be matched in relatively fine-grained ways to the needs of the moment. The trouble is that the major standardized achievement tests we now use are not ideally suited to any but the grouping function. For public accountability, they seek to spread out scores arbitrarily and magnify individual (and thereby social class and ethnic) differences and may therefore mask real gains, especially in low groups. In addition, they are expensive to administer. A much shorter test, perhaps using matrix sampling techniques in which not all children take all items, would suffice for accountability purposes, and some progress in the development of such tests now seems to be on the horizon. For diagnostic purposes, again the item selection techniques of our main standardized tests by and large make them unsuitable for diagnosis. The items are not "clean" enough, and there are often too few on any given skill or topic. Furthermore, giving one big test is not a very efficient way of diagnosing, other than for initial grouping purposes.

4. Is the goal of increasing achievement test scores an appropriate, feasible goal for federal, state, and local educational policy-makers?

The goal of increasing test scores per se is a trivial one and could possibly become a pernicious one. That is, if the goal leads to coaching for the test or to efforts to "beat the test," it can pervert education. The possibility cannot be ignored.

The goals of eradicating illiteracy and of increasing educational achievement generally—so long as these do not lead to an overriding preoccupation with test scores—are surely appropriate goals for the educational system.

The question is: Should such goals be described as local, state, or national? And what are the policy implications of so construing them? Whose responsibility is it to increase educational achievement? We note that none of the "promising directions" described in number 3 above requires federal or state action to be implemented. Of course, their implementation can be improved and facilitated and encouraged by federally funded research and development.

If so, what are the anticipated differential impacts by income and SES groups?

There is the risk that the checking of minimal competencies may only be the measuring of accumulated deficits. The consequences to those who do not pass must be assessed. If success on tests (for purposes of graduation or promotion) is achieved by four-fifths of a suburban school system, but only one-third of a central-city system, the consequences could be serious for domestic tranquility as well as social equity in a world where a high school diploma, regardless of intrinsic meaning, is frequently a ticket to particular jobs.

If social policy increases the trauma associated with educational sorting devices by increasing the stigmas of scholastic failure, more is at stake for the society at large than the elevation of average standards.

5. Present efforts of states to establish "minimum competency standards" inevitably raise questions of appropriate grade levels and cutoff points. What potential problems should be addressed in the states' efforts to institute minimal competency requirements?

The NAEd Panel believes that any setting of statewide minimum competency standards for awarding the high school diploma—however understandable the public clamor which has produced the current movement and expectation—is basically unworkable, exceeds the present measurement arts of the teaching profession, and will create more social problems than it can conceivably solve.

It is basically unworkable because in many populous states cut-off points for a passing grade that are politically and educatively acceptable to parents,

pupils, and educators would have to be so low that an overwhelming majority of students would be allowed to pass. This would make the diploma standard almost meaningless. The recently reported experiences of the Board of Regents in New York State in not approving their own mandated high school graduation test because it was too low is a classic case of what will happen in one form or another in all states. The present measurement arts of educational testing are simply not up to the ambiguous expectations reflected in most state legislation. In a field like English composition, for example, objectives and testable standards are subject to an almost infinite variety of professional disagreements and technical quibbles.

Setting minimum competency standards for high school graduation is the final step in a completely unsuccessful effort to reduce to precise behavioral terms what education is all about. Attempts to impose competency-based teacher education and competency-based preparation of school administrators have surely failed. The effort to determine and assure minimum competency standards for high school graduation will also fall of its own weight, for the scaffolding of existing test designs is too weak to carry such an emotionally laden and ambiguous burden. Continuing any extensive efforts and funds in this direction is wasteful and takes attention away from the major tasks of improving our schools.

As noted elsewhere in this report, the social consequences of failing those in our society who have accumulated educational deficits does not cure their problems. It only exacerbates them.

However, the Panel is in agreement that a series of standardized tests at the lower grade levels used for diagnosing individual student weaknesses, pinpointing remediation needs, and building public pressures if school-wide performances in basic skills continue over time to be consistently low could be positive influences on student learning. Such tests can have negative effects if they become the sole magnet of educational energies. But used with care, and with increased public (particularly teacher and parental) understanding of the diagnostic and therapeutic use (and limitations) of tests, minimum competency testing can be a positive educational development.

"Going public" with state-wide minimum competency tests at lower and medium grade levels is useless unless it stimulates better teaching. Unfortunately, there is not today a situation in which the public seems willing to trust the judgments of teachers and principals in certifying their children's competence. To that extent we have as a nation "bought" the argument of the original test builders that judgments of competence would be more reliable and less biased when based on scientifically constructed tests than when left to the judgments of school people. It is not only the test-builders who believe this. School people themselves probably believe it, and many parents and other lay people are adamant about it. So tests are needed. But what kinds? The term "minimum competency" implies a criterion-referenced rather than a norm-referenced test. But the validity—and social acceptability—of such a test depends on acceptance of the definitions of competence that are built into the tests. How can agreement on these matters be secured?

There seem to be two contradictory routes being followed. One is to use the global standardized tests that now exist. Because they are global there is general agreement that they are testing the right kinds of things. They may work for skills such as reading or computation. They will not work so well as soon as knowledge rather than skill is at stake (for example, in history or science), because there will be many different ideas of what content should be taught. Tests that sample widely across content might be a compromise solution, socially workable, although not very "clean" or attractive to instructional purists.

The second approach has been to ignore the global tests and start building lists of competencies that people think are important for children to acquire. Often these lists are developed by or in conjunction with citizens' committees. The politics of the idea is superficially attractive, but the result is a *reductio ad absurdum* of the basic idea of criterion-referenced standard setting. Knowledge and skill cannot be divided into hundreds of discrete objectives, valid tests for all those separate objectives probably cannot be written, it would take impossible amounts of time to test everyone on all those objectives, and we would not know what to do with the information once we got it.

We are then forced back to the concept of general tests for minimum competency in a very limited number of basic skills. As has been pointed out, using such tests for high school graduation alone is bound to fail. Experimenting with critical "competency points" at selected lower grade levels has substantial merit.

6. There is a growing emphasis on holding the schools responsible for what and how well students learn. What is the most likely administrative level—school building, school district, state, etc.–for producing measurable changes in educational achievement?

The school district or, preferably, the school should be held responsible for what and how well students learn. It is that administrative level which should be held responsible for producing changes in educational achievement. It is important to hold responsible those who actually have the power to do something. The importance and expense of education tempts the public and politicians to support educational panaceas. Instead, we should be insisting that local schools perform.

What role, if any, should the federal government play in this accountability movement?

The federal government should support research on the factors which produce measurable changes in educational achievement and should foster the development and dissemination of information designed to increase public (especially parental) understanding of the uses and abuses of standardized tests.

7. Is there a role for the U.S. government in producing or sanctioning "voluntary national tests," or in monitoring or assuring the quality of tests used by state education agencies and local education agencies? If so, what is that role and how should it be structured and with what safeguards?

There is a broad array of possible federal roles with respect to testing. At one extreme is a single national mandatory test for high school graduation. At the other extreme is providing information. There may be one test or many. Test use may be mandatory or voluntary. Test development may be publicly or privately financed. Public financing may be at the federal, state, or local level. Test development may be under public or private auspices. Quality control may be publicly or privately financed and managed. Test administration may be national, state, or local; financing of test administration may be federal, state, or local. Moreover, the results of the test may be used for making decisions about individual students, classrooms, schools, school districts, states, or the nation. The results may also not be used for making decisions.

A mandatory national test would have the federal government interfering with the responsibility of the states for education. Why should the federal government produce a national test? Is it to indicate what the federal government thinks the schools should be teaching? If the tests are to provide information about the progress of American education, the National Assessment of Educational Progress accomplishes that purpose without infringing on the states' responsibility.

The notion that the federal government should be the Bureau of Standards for educational testing is both professionally unsound and politically dangerous. Since tests reflect the educational goals for many students and teachers, the federal approval or disapproval of tests is in effect having the federal government determine what the schools should teach. It was this fear that led the Executive Committee of the American Association of School Administrators to oppose the National Assessment. This fear was partially overcome by placing the initial responsibility for the National Assessment in the hands of the Education Commission of the States.

While acknowledging the legitimacy of the federal government as a partner in improving educational opportunity and quality, the Academy wishes to underscore the importance of maintaining the principal constitutional responsibility of the states in this important area of public concern.

In short, the Academy strongly recommends against the creation of "national tests"—mandatory or voluntary. Federal approval of tests is also risky. The power to approve tests is ultimately the power to approve the curriculum. The logic is clear: Those who make up tests that confer educational status on individuals determine what is taught in the local classroom. The federal government should not be thrust into the position of fostering a national curriculum by inadvertence.

Federal funding of agencies which develop tests or of agencies which engage in quality control of tests is less problematical. The key is that a variety of

such centers be supported. What is important is to avoid a single, official, federal testing monolith.

And, of course, the federal government can and should support research designed to improve the quality of tests, and programs to increase public understanding about their use.

8. Some scholars challenge the validity of distinctions drawn between "norm-referenced" and "criterion-referenced" tests. What is the significance of this distinction, if any, for public policy?

Challenging the distinctions between the tests

Those who challenge the validity of the distinctions between "norm-referenced" and "criterion-referenced" tests argue that norm-referencing and criterion-referencing are processes that can be applied to tests. Either of these processes can be applied to one and the same test. A test is "criterion-referenced" or "norm-referenced" depending on which of these processes has been applied to it.

Norm-referencing simply means that statistical operations have been performed on test data to make possible the interpretation of the scores in terms of "norms," that is, in terms of the standing or rank of any given score in a distribution of scores that has been determined for some given sample or population of students to whom the test might appropriately be given. For example, a particular score on a reading test might be found to stand very high among all scores made by third-graders in urban schools—say, at the 95th percentile, meaning that it stands above 95 percent of all scores made by students in that population, but below the top 5 percent of all scores of those students. The same test score might be found to stand much lower in some other group—say, at the 60th percentile for a representative sample of fourth-graders in urban schools, or for a sample of third-graders in suburban schools.

Over many years, norm-referencing has been found to be a convenient and generally well-understood way of interpreting test scores. It permits assessing a student's educational progress in comparison to the level or degree of progress that students of a given age or grade typically make. Using test norms, a teacher can readily confirm that a child is making markedly slower progress than other children of the class. A school principal can compare the progress of third-graders in his or her school with that of third-graders across the country. In many circumstances, such comparisons are valid and useful in judging whether efforts to teach or educate students are attaining expectations.

Norm-referencing of tests does have certain drawbacks, however. It depends on the availability of suitable normative data. Good normative data are difficult to collect; publishers of standardized tests often spend large amounts of money in the effort to establish representative and up-to-date norms, but norms can easily become out of date or inaccurate for a number of reasons. Norm-referencing inevitably implies comparison of a student with other students. While these comparisons may be fair when the student is compared to others in the student's own norm group, they may not be completely fair, and they may even become invidious, when students are compared to students in norm groups other than an appropriate one. For example, it might be regarded as unfair to compare urban students' reading scores with those of students in suburban schools. (On the other hand, such a comparison might be thought by some people to be very useful.) The major drawback of norm-referencing, however, is that it does not give concrete information on exactly what level of ability, knowledge, skill, or performance is represented by a given test score, even if that test score is a highly accurate measure. Thus, while it may be informative to know that a third-grade student attains the 60th percentile among third-graders on a standardized reading test, this does not tell how well the student can read; i.e., what level of text difficulty and complexity he or she can handle. This kind of information would be forthcoming, if at all, only if the test is also criterion-referenced.

Criterion-referencing means that the scores on a test can be interpreted in terms of concrete, specifiable, and describable levels of ability, achievement, or performance. That is, a criterion-referenced test score is one that can be interpreted as indicating that the student possesses some specified level of knowledge, understanding, or skill. Sometimes this means going into minute details as to exactly what the student knows or does not know—or can or cannot do. Scores on an arithmetic test, for example, might be "criterion-referenced" in the sense that

they would indicate exactly what addition facts the student has failed to master. On the other hand, a much more generalized idea of what a test score means is frequently desired. It would be too much to expect a reading test score to indicate exactly what words a student could recognize; there are simply too many words to be tested, and even if one could develop a test that could be "criterion-referenced" in this way, the information might be almost overwhelming for a teacher. Rather, it would be desirable to "criterion-reference" a reading test in terms of the representative levels of text or passage difficulty that the student getting a given score could be expected to comprehend. If one can decide what level of text difficulty represents "functional literacy," the test could be said to yield information on whether a given student has attained functional literacy.

From the above, one can see not only that it is possible either to "norm-reference" a test or to "criterion-reference" it, but also that it is possible for a test to be both "norm-referenced" and "criterion-referenced," depending on what kinds of research operations have been performed in connection with the test.

In principle, norm-referencing is comparatively simple, even though it may require large and costly research operations. To "norm-reference" a test, it is administered to large and representative samples of one or more defined groups that the test is likely to be given to, or is appropriate for—e.g., third-graders, fourth-graders, and fifth-graders, or to urban third-graders, suburban third-graders, rural third-graders, etc. Score distributions are then set out in each of the groups so that the test user can find the level of a given score with reference to an appropriate group. These score distributions are usually presented in "tables of norms," and there are various systems offered for interpreting scores—percentiles, "standard scores," "T-scores," and the like.

Criterion-referencing a test is in many ways much more difficult than norm-referencing it. It requires a detailed analysis of what kinds and amounts of knowledge are represented by a given score level on the test. Obviously, a very high score would indicate that the student has mastered all or nearly all the knowledge or skills tapped by the test, and a very low score would indicate that the student has failed to master all or nearly all of the knowledge and skills. Even so, interpreting these scores entails communicating exactly what knowledge and skills are or are not measured by the test, and this is often not easy to do. Most good tests are designed on the basis of a detailed specification of the abilities, knowledge, or skills they are intended to measure, but this does not always guarantee that such abilities, knowledge, or skills are in fact measured by the test. Even with painstaking care on the part of the test maker, scores sometimes reflect abilities and achievements (or lack thereof) that were not intended to be measured. For example, a test designed to assess knowledge of the grammar of a foreign language might turn out to be much more a measure of knowledge of the vocabulary of that language.

Even though it might be relatively simple to interpret very high and very low scores on a test, it is harder to "criterion-reference" scores in the middle ranges of the distribution. Sometimes it is possible to arrange the skills and knowledge tapped by a test in order of relative difficulty, and then to establish a more or less close correspondence between the test scores and the ranking, in terms of relative difficulty, of the skills and knowledge. A medium-level score, then, might be found to mean that the student getting such a score has mastered the easier skills, but has failed to master the more difficult skills. Thus, a medium-level score on a spelling test might mean that the student can spell correctly all or nearly all of a list of easily spelled words, but would fail to spell correctly many words that are often misspelled. Note that in this case criterion-referencing would entail giving detailed lists of words that students at various score levels could or could not spell.

This kind of criterion-referencing is generally unfeasible for the traditional types of standardized achievement tests in the various school subjects—reading, arithmetic, science, social studies, etc. For one thing, these tests attempt only to include small samples of all the skills and knowledge that are ordinarily taught in these subjects. It would be difficult to put scores on such tests in correspondence to highly detailed lists of specific skills and knowledge that students at different score levels would possess. And even if this were done, the results would be so voluminous that it might be burdensome for teachers to utilize them effectively. Only when criterion-referencing of achievement tests can use rather generalized descriptions of skill and

knowledge levels is it possible to "criterion-reference" most of the standardized tests commonly used in today's schools. Such criterion-referencing has not generally been attempted, but it is quite feasible and certainly desirable. For example, a standardized test of competence in a foreign language can be criterion-referenced in terms of generalized levels of proficiency, e.g., "able to read simple prose," "able to carry on a conversation on a general topic with a native speaker of the language," etc. Similar generalized descriptions could be given for ranges of score levels on achievement tests in other school subjects.

Because of what are perceived as difficulties in criterion-referencing of traditional types of tests in the regular school subjects, some educational measurement specialists have recommended the development of tests that would lend themselves readily to criterion-referencing. Generally, this has meant the development of a large number of very brief tests—even individual items or "exercises"—that are specifically targeted on measurement of particular educational objectives. For example, one might have a test to determine whether a student has mastered a particular skill (e.g., the operation of "carrying" in adding a column of numbers) or understands a particular concept (e.g., "the distribution of labor"). In this sense, and only in this sense, one can say that there are certain types of tests that are specifically "criterion-referenced" and not "norm-referenced." Yet, one could also, as it were, "norm-reference" these tests by finding out what percentages of various groups can pass them. (A good illustration of a type of "norm-referencing" of these highly targeted tests is afforded by the sample results published by the National Assessment of Educational Progress on what percentages of different groups, by age, grade, and region of the country, can pass particular exercises included in the NAEP assessments.)

Highly targeted "criterion-referenced" tests of specific skills and achievements are more useful in the day-to-day conduct of teaching than they are in the overall evaluation of students' progress at the end of the school year or at the end of a major segment of the student's educational career. Standardized general achievement tests are usually more suitable for the latter purpose.

Supporting the distinctions between the tests

Those who argue that norm-referenced tests are different from criterion-referenced tests stress that distinctions between the tests arise from the selection of behavior to be sampled, the definition of the universe of the test. This defined universe is the "criterion." If the criterion is clearly defined and the universe properly sampled a trained user of the test can judge more rationally its usefulness for the purpose he has in mind than is possible with a norm-referenced test. As an example, the lack of definition of the universe sampled by the SAT made it impossible for the Review Panel to judge the significance of the score decline in any terms other than self-contained statistical ones.

Implications for public policy

It would be a mistake to regard "criterion-referenced" tests as necessarily better, more desirable, or more useful than "norm-referenced" tests. It would be wrong to decree that all "norm-referenced" tests be discarded or outlawed and that only "criterion-referenced" tests be permitted. One and the same test can be both "referenced to criteria" and "referenced to norms." Further, both "norm-referenced" and "criterion-referenced" tests have their uses in education. And both norm-referencing and criterion-referencing of tests have distinct and relevant functions in education.

Recommendations from the Education Commission of the States

The Education Commission of the States has spent a year studying the minimum competency testing movement. It has produced several monographs, kept a record of state deliberations and legislation (see *Back-to-Basics*, NSBA Research Report #1978-1, pp. 17-20, for such a compilation), and on the basis of its studies, adopted policies which it recommends to state and local governments.

ECS notes, in the introduction to its policy, that minimum competency testing is a complex phenomenon, more so than its name would suggest. It finds many and conflicting motivations among those who advocate such testing and in the purposes intended by different states. Also, ECS notes that there are several technical considerations about testing that cannot be ignored.

"The laws and regulations that have been enacted provide evidence that some programs are designed to evaluate the performance of schools and others the performance of students. Some plans are designed to provide accountability and to save money, while others are directed primarily toward the improvement of instruction and remediation." Technical questions include "issues related to format, test conditions, validity (How well does the test measure what it was designed to measure?), and reliability (How consistently does the test measure that?)." Paper and pencil multiple-choice formats may not measure what are regarded as competencies, but other kinds of performance tests are time-consuming and expensive to administer. Also, the establishment of cut-off points and concern about cultural bias are political as well as technical issues.

The policy recommendations below, adopted June 17, are certain to be influential throughout the nation in the coming months.

Policy Issues

A. *Discussion of Issues*

Responsible policy at any level will involve comprehensive and cooperative study by educational and political leaders of the political and technical realities as part of the search for commonly acceptable solutions for improving basic skills. The primary thrust must be for policy which supports improvements in the teaching/learning process. A test by itself is not a solution to learning problems. Effective programs must include attention to the instruction that precedes and follows the test. The goal of improving student achievement should be a primary focus of discussions of testing and standard setting.

To encourage a positive and permanent impact on the instruction of basic skills, policymakers should avoid the tendency of other accountability-laden proposals to respond to public criticism and demands for change with a mechanical solution that has little or no relationship to how students learn.

The decision of whether or not to implement a minimum competency testing program should involve a weighing of the positive and negative consequences of either decision. The following list* of possible outcomes can serve as a starting point for analyzing whether or not minimum competency testing would be a good approach to improve the attainment of basic skills.

Possible Positive Outcomes:

- The question, "What is a high school education?" must be squarely faced.

* This list is from *Competency Tests and Graduation Requirements*, National Association of Secondary School Principals, Reston, Va. 20091, 1976.

- The statements required for each course will likely result in carefully organized teaching and carefully designed sequential learning.
- Slow learners and underachievers will likely receive direct and immediate attention.
- Courses of study will likely be revised to correct identified deficiencies.
- Subjects leading to the development of competencies will receive additional emphasis.
- Alternatives and options not requiring attendance in class will likely be broadened.
- The senior year may gain more holding power because of a new focus on requirements and options.
- The community will know the minimum performance required in specific subject areas for the diploma.

Possible Negative Outcomes:

- Confusion over the meaning of high school diploma will continue if each district identifies its own level of competencies and performance indicators.
- The emphasis on pragmatic and practical competencies may result in erosion of liberal education.
- The emphasis on measurable outcomes could result in less attention to outcomes that are difficult to measure.
- The record-keeping system could become burdensome to teachers and administrators.
- The conflict between "humaneness" and "accountability" may be intensified as criteria are established and clarified.
- Community disagreement may arise over the nature and difficulty of competencies.
- Dropouts could increase depending on the level of the minimum competencies.

The complexity of issues and range of unanswered questions must be directly and deliberately addressed. Policy should encourage research and the incorporation of findings into any decision to move toward minimum competency testing. If the conclusions of research papers completed by scholar/practitioners for the series of conferences sponsored by ECS and NIE (National Institute of Education) are any indication, there are many concerns that should cause people to move carefully toward program design and implementation. States should examine the differences that existing programs are making when embarking on a plan to improve achievement.

B. ECS Policy Positions

ECS has adopted the following policies in regard to minimum competency standards and testing for use by education policymakers at federal, state, and local levels. As new knowledge is available from research and experience, these policies may be amended.

1. ECS supports state efforts to improve educational achievement and to increase the educational attainment of underachieving students to levels which will enable them to participate effectively in society, including obtaining and retaining a job. Setting standards for graduation, grade-to-grade promotion, or placement in programs and specifying methods for determining when those standards have been met are common practices in education. Some standards are set by the state, some by the local school district or school. Teacher-made tests, department-wide examinations, standardized tests, minimum competency tests, teacher judgment, and demonstration of mastery by student performance are some of the methods used to determine when pupils have met the standards. Use of minimum standards and minimum competency tests to determine when the standards have been met is one option. There is not, however, a body of knowledge or research data to support a conclusion that this method is better than others. Therefore, ECS advocates that states and local districts adopt a comprehensive plan for setting standards and establishing criteria for determining when those standards have been met. Such a plan should avoid the selection of standards that are too narrow and the reliance on a single assessment technique.

   Important Considerations:

   - Minimun competency testing and standards are one part of a larger national concern about achievement in education. Among those dedicated to and actively working for improvement in learning and achievement in the basic skills, there are strongly held, often widely differing views as to the promise that minimum competency testing holds for attainment of these objectives.
   - A variety of methods is presently available for assessing students, determining standards, and establishing quality controls within curricular programs. Leaders at both the local

and state levels are exploring other options in addition to competency testing for improving student achievement. The diverse approaches include use of state-wide assessment data to improve programs, management by objectives to evaluate instruction, revision of high school graduation requirements, individualized learning programs, and targeting of individual learning problems for specialized help. The decision not to set minimum standards through technical and political issues that must be addressed when designing and implementing a minimum competency testing program.

2. In any consideration of the establishment of minimum competency testing and standards, the purpose for the testing and standards should be clearly determined.

Important Considerations:

- The primary reason that minimum competency testing and standards are being implemented is to ensure a minimum level of achievement by pupils. In designing minimum competency legislation or policies, special attention should be given to see that this purpose will be achieved. For the test results to be helpful, it will be necessary to understand the limitation of and uses for tests.
- Minimum competency standards must not be confused with the goals and objectives of education for the public schools. Goals and objectives should be designed to foster maximum development of each pupil. Minimum standards are only one aspect of the total education program. The purpose of minimum competency standards is to set a lower limit of acceptable achievement in designated basic skills for pupils. Minimum standards are not adequate measures of the output or effectiveness of individual teachers, the adequacy of a total school program, or the full range of an individual's skills and abilities.
- One promising use of minimum competency testing for ensuring achievement in the basic skills is to use these tests as an early warning device or for diagnostic purposes early in the elementary grades. Low-achieving pupils should then be placed in instructional programs designed to improve their achievement in those skill areas in need of improvement. Children should be taught those skills that the tests indicate they do not know, not recycled through a program that does not effectively meet the individual needs.
- The use of a single test of competency to determine student qualifications for graduation or grade-to-grade promotion is an inadequate measure. Given the limitations of a single test and the importance of such decisions to the affected individual, other evidence such as the teachers' observations and judgment should play a part in these decisions.
- Since the source of many of the problems related to low achievement has family and societal origins, the school will not be able to solve the problems without the help of the family and community.* A partnership that defines the role of school, home, and community should be the basis of a minimum competency program.

3. Diagnosis of learning problems of pupils who do not attain minimum standards and implementation of instruction programs that focus on their particular deficiencies should be components of any minimum competency program.

Important Considerations:

- The enormous range of differences in individual abilities must be recognized. There will always be some who fall below the minimum standards in spite of the best efforts of the school and student. Minimum competency laws or policies should recognize individual differences and make provisions which are educationally productive for those students who initially cannot pass the test.
- Frequently, minimum competency policies call for retention or remediation for pupils who score below the minimum. The evidence from research is that routine repetition of a grade is not an effective means for improving achievement. Remediation in the basic skills is sometimes effective and sometimes not, often because the rates and modes in which children learn may not be clearly identified. Remediation is more likely to be effective if

* A blue ribbon panel of experts headed by Willard Wirtz and commissioned by the College Entrance Examination Board (CEEB) found lower standards to be only one of the many explanations for the decline in test scores. Other causes included societal changes, large numbers of hours spent watching television, and increasing numbers of students living in unstable family situations or working at full- or part-time jobs while attending school.

started in the early grades and if the instruction is targeted at the particular learning level of the individual. Therefore, staff expertise and funds should be provided which will make it possible for schools to design and implement remedial programs that are based on individual needs and learning styles.

4. The roles of all levels of government regarding minimum competency testing should be clearly delineated. The federal government should not become involved in minimum competency testing in states or establish a national testing program or national minimum competency standards. Technical assistance to the states, support for research and development, and data collection on a sampling basis and reporting of national levels of knowledge are important functions that the federal government should support. Generally, the state should establish the broad goals and objectives for the state, set the limits within which local districts must operate to achieve those goals and objectives, provide funds and technical assistance to districts in program development and staff training, and give the local districts and the schools latitude to design and implement educational programs to achieve those goals and objectives. Likewise, if a statewide minimum competency testing program is implemented, states should provide for local districts to have an important role in establishing standards and tests.

   Important Considerations:

   - The states, with constitutional sanction, have the primary responsibility for public education.
   - A federal role in minimum competency testing and standards might include technical assistance to states, development of benchmark data on changes in achievement by population samples in important skills and competencies, and support for research and development of better methods for competency assessment.
   - Standard setting is a shared responsibility by the state and local education agencies. A strong rationale exists for making many educational decisions at the district or local school level within each state. It will be important to determine the level at which decisions should be made about types of competencies to be identified, desired level of achievement, amount and types of remedial assistance, types of data to be collected, and measures of achievement.

5. The procedures and processes for developing and implementing standards should encourage a broad base of active involvement among parents, students, educators, and other concerned citizens.

   Important Considerations:

   - Deciding what to include under basic skills, determining what the standards should be, developing good tests, and designing effective instruction programs are complex problems for which standard solutions do not exist. Because educators do not have all the answers, research directed toward the identification of better approaches to solving learning problems should be an integral part of the decision to test.
   - If minimum competency testing and standards are to be implemented, sufficient time should be provided for planning and development work and for teachers and students to adjust to the new requirements. Components of the program which must be built into the time line include the selection of competencies, evaluation of the present curriculum and possible design of new instructional components, development and piloting of tests to establish reliability and validity, and the development of a diagnostic capability.
   - While a wide range of useful skills has been suggested as appropriate competencies for schools to address, selection among basic or life skills, process and social skills, or skills and knowledge within each subject area can be made within the identified jurisdiction. Because competency testing may or may not be used to measure student progress in all areas, it will be necessary to determine how large a role the minimum standards should have in the definition of a quality education.
   - In selecting tests to be used, whether done at the state or local level, special attention should be given to determine that the tests measure what the curriculum is designed to deliver. This is often more difficult to do when tests are selected and mandated by the state, because of differences in instruc-

tional programs among school districts within a state.

6. If minimum competency testing and standards are to be used as a factor in promotion or graduation decisions, careful attention must be given not only to the content and administration of the test, but also to the provision of compensatory and bilingual education programs, remediation, and any other instructional strategies necessary so that application of the standards will not discriminate against minorities or the disadvantaged.

   Important Considerations:

   - The demands for higher educational standards are bound to challenge many of the changes made in schools during the past decade to promote equality of opportunity. During that time, modifications in curriculum and standards were made in an effort to hold students in school. A decline in the dropout rate to an all-time low of 25 percent in 1977 is evidence that the changes may have encouraged a larger number of previously disinterested students to remain in school.
   - The College Entrance Examination Board panel found two-thirds to three-quarters of the decline in the SAT scores between 1963 and 1970 was related to the "notable extension and expansion of opportunity in the United States." Policymakers should study the impact that standards may have on students who may choose to drop out of school because they are unwilling or unable to meet the established standard.
   - Minority groups are urging consideration of the impact of testing on culturally and linguistically different populations. In statewide tests in some states, minority students have fallen below standard in large numbers. The needs of such students should be addressed without limiting access to opportunity through tracking or denial of the high school diploma. Alternative methodologies or environments will need development and support to meet the needs of students not achieving in the traditional setting.
   - Other special need groups include special education students who may lack the capacity to pass a proficiency test, handicapped students who need special forms of the test, gifted students who deserve challenging work not provided by an emphasis on the minimum, and highly mobile students who must phase in and out of various local- or state-determined requirements.

7. Development of minimum competency tests and administration of minimum competency standards cost money. Sufficient funds should be provided to do the development and implementation work well. Funds are also necessary to implement good compensatory education, bilingual education, and remedial or competency-based education programs and for staff training if improved achievement in the basic skills for all pupils is to be attained. These requirements will likely call for revisions in priorities for use of existing funds as well as provision of additional funds.

   Important Considerations:

   - Funding decisions should be made as an integral component of minimum competency policy. Designing a testing program that directly affects students should not be undertaken without determining how the program, and the necessary remedial assistance to students, is to be financed. Some of the costs to be considered are:
     a. Administrative costs—record keeping and reporting expenses.
     b. Enforcement costs—staff to monitor, evaluate, and police.
     c. Costs in lost support to programs previously receiving funding.
     d. Test development costs.
     e. Test administration & interpretation.
     f. Remedial work.
   - Policymakers need to consider which are the most cost effective ways of assisting students who do not achieve at minimal levels. Because it will cost more to raise some students to the minimum than others, districts with large numbers of hard to educate students should be provided the necessary funding to implement remedial programs. Funding alternatives to consider include categorical grants based on number and type of low achievers, indexing the basic financing formula, and reprioritizing local budgets to create the necessary funds for remediation.
   - Finances will necessarily determine how

idealistic or realistic a state can be in setting standards. In the idealistic sense, "If one believes that everyone can master most things given enough time (and time is money), then it follows that the more people who must master a skill, the more it will cost society."* For example, the resources needed to raise the reading level of a student from second to eighth grade level will be significantly higher than for a single semester of remedial reading that is not aimed at a specific criterion. The resources and priorities of each state and locality should be evaluated when determining how high the standard should be.

8. Minimum competency legislation and/or policies should have as their primary objective the improvement of pupils' mastery over the basic skills. Since universities and colleges can make a significant contribution to the attainment of this objective, policy makers should give attention to ways in which the expertise located in institutions of higher education can be brought to bear in implementing a minimum competency program.

   Important Considerations:

   - Teacher training institutions have a significant impact on the preparation and selection of persons who teach in the nation's schools. It is important for these institutions, through appropriate pre-service and in-service programs designed in cooperation with SEA's and LEA's, to insure that teachers, principals, and all other school personnel are adequately trained to carry out their responsibilities.
   - The state of the art in the development and administration of tests leaves much to be desired. Universities and colleges should perform research and development functions in furtherance of knowledge and practice concerning the technology of testing.
   - If minimum competency standards and testing programs are to ultimately result in improvement of students' educational achievement, additional research focused at the teaching-learning process is required. Too little is known about the identification and treatment of specific learning disorders, potential remediation techniques, and, more generally, what factors make educational programs effective. Until more is known about the learning process, our inability to teach some pupils the basic skills is likely to continue.

* Barry Anderson, *Costs of Minimal Competency Testing*, a paper prepared for the ECS/NIE conferences on minimum competency testing.

The NSBA "Minimal Competency Testing" Survey

When Joseph Califano, Secretary of the Department of Health, Education, and Welfare, invited NSBA to send a delegation of local school board members to the National Conference on Achievement Testing and Basic Skills (held in March 1978), NSBA Research conducted a survey among shool board presidents in order to better inform the NSBA delegation. The survey attempted to determine the thinking of school board members about some of the issues involved in competency testing.

The Sample

The survey was mailed in mid-January to more than 3,000 school board presidents, and 878 usable questionnaires were returned within 21 days, the deadline for data processing. Unfortunately, the short time between the announcement of the conference and the conference itself made it impossible to send a follow-up mailing of the survey. More than 200 complete questionnaires arrived too late.

Of the 878 persons who comprised the total sample, 79 percent were board presidents and 8 percent ex-presidents; 13 percent were superintendents or other high-ranking administrators who completed a survey in lieu of its intended recipient. On most issues, board and administrator respondents concur with each other; where they do not, the fact is noted. All 878 responses are reported here, representing 878 school districts, or 5.5 percent of all the nation's public school districts.

The sample is a sensitive and informative one, representing the diversity of school districts by size, geography, and type of school district. However, it is not a perfect cross-section of the school board population, because it is weighted toward responses from districts of more than 1,000 enrollment (see Table I).

Most Pressing Concerns

One always is concerned that respondents to a survey may be only those persons with a zealous interest in the topic, in this case, minimal competency testing. To find out whether this was so, the survey asked, "From the following list, (of 17 items) please check the THREE most pressing concerns of your school board in your district." Table II shows the responses, in rank order. One respondent in three cited "improving the quality of teaching" and one in ten cited "declining test scores." Among the 125 write-in items, the majority were about money, inflation, balanced budget, finance, decreased state aid, resistance to property taxes, and the like. The term most often used was "mandates" as in "state mandates," "mandated programs without full funding," "fitting mandated programs into budget," or "inadequate state funding coupled with mandates."

The survey respondents are certainly not obsessed with test scores but rather reflect a diversity of concerns. Thirty-seven percent of board members cite the quality of teaching, compared with 26 percent of superintendents. Thirty-nine percent of superintendents cite collective bargaining with teachers, compared with 32 percent of board members.*

* Some other variations show that energy and inadequate facilities are more a concern of small districts than of larger districts; declining test scores is more worrisome in larger dis-

TABLE I: THE SAMPLE

Region

East. Approximately 23% of all board members reside here, in Maine, New Hampshire, Rhode Island, Connecticut, Vermont, Massachusetts, New York, New Jersey, Pennsylvania, West Virginia, Delaware, Maryland, and the District of Columbia. 18% of the sample is from the East.

Midwest. Approximately 40% of all board members reside here, in Ohio, Indiana, Illinois, Michigan, Minnesota, Wisconsin, Iowa, North Dakota, South Dakota, Kansas, Nebraska, and Missouri. 42% of the sample is from the Midwest.

South. Approximately 22% of all board members reside here, in Kentucky, Tennessee, Virginia, North Carolina, South Carolina, Georgia, Florida, Alabama, Mississippi, Texas, Arkansas, Oklahoma, and Louisiana. 21% of the sample is from the South.

West. Approximately 15% of all board members reside here, in Arizona, New Mexico, Colorado, Nevada, Montana, Idaho, Wyoming, Utah, California, Washington, Oregon, Alaska, and Hawaii. 18% of the sample is from the West.

Enrollment

Under 1000. Approximately 50% of all school board members serve in these districts. 7% of public school children attend these districts. 25% of the sample is from these districts.

1000—4999. Approximately 38% of all school board members serve in these districts. 29% of public school children attend these districts. 34% of the sample is from these districts.

5000—9999. Approximately 8% of school board members serve in these districts. 18% of public school children attend these districts. 15% of the sample is from districts of this size.

10,000—24,999. Approximately 4% of school board members serve in these districts. 18% of public school children attend these districts. 19% of the sample is from districts of this size.

25,000+. Approximately 1% of all school board members serve in these districts. 28% of public school children attend these districts. 7% of the sample is from districts of this size.

Type of District

Predominantly *rural*: 30% of respondents.

Predominantly *small town*: 27% of respondents.

Predominantly *suburban*: 31% of respondents.

Predominantly *urban*: 12% of respondents.

Minimal Competency Testing: Issues & Attitudes

Familiarity with the "Meaning" of the Concept

Most respondents said they were either "very familiar" (39 percent) or "somewhat familiar" (55 percent) with the concept of minimal competency testing. Only 5 percent said they were not familiar with it.

As a point of departure, and particularly for those who were not familiar with minimal competency testing, the survey form provided a "definition," that is, "testing programs which attempt to learn whether each student is at least 'minimally competent' by the time that student graduates from public school. There are several vital issues about this kind of testing: Should there be such testing programs? Who should design or run them? What should their content be? What uses should be made of the scores?"

Opinions and Attitudes

The survey then asked respondents to express agreement or disagreement with 11 issues about competency testing. Responses are shown in Table III.

Most school leaders seem to be sympathetic to the premises that underlie the competency testing movement: 76 percent agree or tend to agree that "every child in public schools should be required to master certain skills"; 93 percent agree or tend to agree that "a high school diploma

tricts than in small districts. Declining enrollment is more a problem for suburban and urban districts than for small town and rural districts. Improving teacher quality seems most pressing to respondents from South and West regions; collective bargaining and declining enrollments are of relatively little concern in the South, but energy costs and inadequate facilities rank high in southern concern. Cutting programs and cutting staff, declining enrollments, and declining tax base are of more concern in East and Midwest regions than in the West and South.

TABLE II: MOST PRESSING CONCERNS

The issues on the minds of board members and administrators are, in order:

Improving the quality of teaching	36%
Collective bargaining with teachers	33%
Declining enrollments	28%
Cutting programs to balance budget	27%
Educating handicapped students	19%
Federal interference	19%
High cost of energy	18%
Inadequate facilities	17%
Curriculum reform	16%
Steady or declining tax base	13%
Cutting staff to balance budget	13%
Public apathy	12%
Declining test scores	10%
Discipline	10%
Evaluation administrators	6%
Vandalism	3%
Drug abuse	3%

Respondents were asked to select the *three* most pressing concerns in their districts.

TABLE III: ATTITUDES

	Agree %	*Tend to Agree* %	*Tend to Disagree* %	*Disagree* %
A. Every child in public schools should be required to master certain skills.	46	30	14	10
B. A high school diploma should mean that a student has mastered certain skills.	62	31	4	2
C. Every student should be required to pass a minimal competency test in order to receive a high school diploma.	42	34	15	10
D. Giving minimal competency tests before graduation will go a long way toward improving teaching in our schools.	22	40	26	12
E. Most people would agree about which "minimal competencies" each student should master.	8	31	39	22
F. If a student fails a minimal competency test, this means that the *school* has not fulfilled its responsibility.	6	25	39	30
G. If a student fails a minimal competency test, this means the *student* has not fulfilled his(her) responsibility.	11	44	29	16
H. Handicapped students should achieve the same "minimal competency" as all other students.	5	16	34	45
I. Non-English-speaking students should achieve the same "minimal competency" as other students.	20	40	24	17
J. The benefits of minimal competency testing justify the time and dollars expended.	18	44	25	13
K. The time and dollars that would go into minimal competency testing could be better used in actual teaching.	16	27	43	14

should mean that a student has mastered certain skills"; 76 percent say that "every student should be required to pass a minimal competency test in order to receive a high school diploma."

But there is less consensus about the purpose or value of such tests. A lesser 62 percent agree that such testing will have the effect of improving teaching. While 31 percent think student failure on such tests "means that the *school* has not fulfilled its responsibilities," 55 percent think such failure means "the *student* has not fulfilled his/her responsibility." Roughly three out of five respondents (62 percent) think that time and dollars spent on such tests are justified, while roughly two out of five (43 percent) think time and money "could be better used in actual teaching."

Respondents are also not certain whether the same "minimum" should apply to all students. Only 21 percent think "handicapped students should achieve the same 'minimal competency' as all other students," and 60 percent say that non-English-speaking students should achieve the same minimum as others.

Only items A, B, C, and H elicited strong "agree" or "disagree" responses; all other items elicited mostly "tend to" agree or disagree responses. This suggests that respondents do not have strong opinions on these matters. It is reasonable to infer that board members have diverse, contradictory, and perhaps confused thoughts about many of these matters.

Test Failure

School leaders differ in their judgment about how many (if any) students can fail a competency test before there are political repercussions. In general, one respondent in five (19 percent) thinks it desirable to pass every student; this is more true among board members (20 percent) than superintendents (13 percent), and in suburban (26 percent) and Eastern (30 percent) districts than elsewhere.

Thirty-seven percent would prefer that no more than 5 percent fail a test. This, plus the 19 percent who would fail no students, means that more than half of school leaders are politically reluctant to give tests which will fail many students. Thirty-two percent said that at least 85 percent must pass; 13 percent said at least 75 percent must pass. A few persons took offense at this question, calling it "stupid" or making such comments as, "You don't devise such tests so some percentage passes. The test must measure competency. I don't know how many would fail—in some schools, possibly all would fail."

Presumably, test failure will prompt some action. Table IV contains three hypothetical circumstances posed. Answers were consistent with the responses in Table III, showing that respondents would be somewhat more inclined to take action against students than against the staff when students fail in modest (i.e., 7 percent) numbers. Criticism of staff is likely to multiply when students fail in more substantial numbers (37 percent). This suggests the possibility that school leaders might use such tests to evaluate staff as well as students.

As was also reflected in Table III, there is some uncertainty as to whether all students should be held to the same "minimum," as reflected in the responses to the question about a higher rate of failure by minority students than by other students.

Whether to re-teach (coach, remediate) and to re-test is also at issue in these responses. It is interesting that in the event of student failure school leaders feel more inclined to criticize students, staff, or curriculum than to criticize, eliminate, or change the tests.

Only a small portion of respondents would "make the test 'easier' " or "stop giving such tests." Of the persons who wrote in comments, about 20 percent said they would "re-evaluate the test," "check the validity of the test," "re-examine the test; standards may be too high," and the like. If minorities fail in large numbers, about 10 percent of respondents' written comments suggest the test should be examined for bias.

A large portion of respondents, rather, would "work to change the curriculum." Some wrote marginal comments like, "If this were to happen I think the system is in really bad shape," and "Close the school! It is a failure."

Among scholars, education writers, educators, and other school people, there is considerable controversy surrounding the use of minimal compe-

TABLE IV: TEST FAILURE

Suppose that a minimal competency test for high school graduation were given in your area this year, and 7% of the students failed the test. What do you think should be done (Check all *appropriate items):*

	%
Coach the 7% and let them take the test again.	63
Deny high school diplomas to the 7%.	33
Criticize the teachers for letting this happen.	6
Criticize the superintendent for letting this happen.	5
Work to change the curriculum.	47
Make the test "easier" for next year.	1
Stop giving such tests.	4
Other (please explain): _____	26

Suppose that a minimal competency test for high school graduation were given in your area this year, and 37% of the students failed the test. What do you think should be done (check all *appropriate items):*

	%
Coach the 37% and let them take the test again.	50
Deny high school diplomas to the 37%.	22
Criticize the teachers for letting this happen.	23
Criticize the superintendent for letting this happen.	23
Work to change the curriculum.	66
Make the test "easier" for next year.	7
Stop giving such tests.	6
Other (please explain): _____	29

Suppose that a minimal competency test for high school graduation were given in your area this year, and 7% of the white kids failed but 37% of minority students failed. What do you think should be done (check all *appropriate items):*

	%
Coach all who failed and let them take the test again.	56
Deny high school diplomas to all who failed.	21
Criticize the teachers for letting this happen.	15
Criticize the superintendent for letting this happen.	16
Work to change the curriculum.	59
Make the test "easier" for next year.	2
Stop giving such tests.	7
Other (please explain): _____	31

tency testing (and standardized testing of all kinds) for educational purposes. Vehement criticisms of the tests themselves have been more and more frequent in recent months, in light of stepped-up activity in minimal competency testing at the state level. From NSBA's survey, it may be fair to infer that school board members and superintendents do not share the views of those who severely criticize these and other tests.

Competencies

Table III (item E) showed that relatively few board members agree or tend to agree (39 percent) that "most people would agree about which 'minimal competencies' each student should master." This observation leads to two questions: What competencies should be tested? Who should decide what competencies should be tested?

School leaders do not themselves agree on more than a narrow set of examples of what constitutes minimal competency. Given a list of 20 possible items ranging from "basic arithmetic problems" to "reciting the Ten Commandments," at least one person in twelve selected each item on the list. Responses are displayed in Table V.

Reading and arithmetic ranked very high, and writing ranked substantially high. The orientation was most frequently elementary and utilitarian, but often not so. Basic arithmetic problems had 99 percent support, while completing a tax form only 34 percent; reading written instructions had 97 percent support, while reading a novel 39 percent; writing a message had 87 percent support, while a business letter had 69 percent and a poem 13 percent.

Approximately 200 respondents offered 400 suggestions for other items that should be in-

TABLE V: COMPETENCIES

Which of the following do you think should be included in any test of minimal competencies? (check all that apply)

	%
Basic arithmetic problems	99
Following written instructions	97
Following oral instructions	89
Reading a daily newspaper	88
Writing down a telephone message	87
Giving oral instructions others can follow	82
Reading a road map	80
Balancing a checkbook	80
Writing a business letter	69
Discussing the plot of a written story	44
Reciting rules of the road	44
Reading a novel	39
Drawing a map	35
Correctly completing an income tax form	34
Driving a car	31
Being compassionate	29
Discussing the plot of a TV story	27
Explaining the difference between political candidates	24
Writing a 4-line rhyming poem	13
Reciting the Ten Commandments	8
Other	18

cluded in any test of minimal competencies. These diverse responses might be categorized into five groups, here listed in order of frequency: writing and language skills (e.g., basic reading, punctuation, spelling, handwriting, writing a coherent paragraph); math, especially practical math and related skills such as consumer arithmetic; practical knowledge (e.g., swimming, hygiene, nutrition, birth control, marriage and family); job seeking (e.g., interviewing, letter writing, resume preparation, work ethic); and government and citizenship concerns (e.g., history, U.S. government, civics, patriotism). Knowledge of art and literature was mentioned only once.

Interesting but modest differences in priorities do appear in the data.

- Smaller district respondents put a greater premium on giving and following oral instructions, on balancing a checkbook, and on being compassionate than do respondents from larger districts. Larger districts put greater premium on reading a daily newspaper, correctly completing an income tax form, and explaining the difference between political candidates.
- Few regional distinctions appear in the data. Completing an income tax form is much less frequently cited in the South and more frequently in the West. Driving a car is much less frequently cited in the East.
- Board members value reading a novel and being compassionate more highly than do superintendents. Superintendents value reading a road map, writing a business letter, reciting rules of the road, and completing an income tax form more highly than do board members.

Decision Making

Among school leaders the survey shows a marked preference for decisions made locally. There is also a preference for citizen involvement in decisions, but a distrust of legislatures.
Two questions were asked:

> "In different states and locales, different agencies are deciding what 'minimal competencies' should be taught and tested. Of all the agencies below, which *one* would you *most trust* to make such a decision? Which *one* would you *least trust*?"

Responses are displayed in Table VI.

Generally, the responses were uniform across categories of respondents. There were a few exceptions:

- Local boards are most trusted in larger, suburban, urban, and western districts.

TABLE VI: DECISION MAKING

Here are the agencies, listed from most to least trusted:*

	Most %	Least %
Local board of education	17	—
Local committees of teachers	15	1
Curriculum coordinators in a district	8	—
State board of education	6	1
A state-level citizens' committee	6	1
Local committees of principals	5	—
The local superintendent	4	—
An ad hoc citizens' committee in a district	7	5
Each school principal	1	—
The state superintendent	1	—
Subject matter experts at universities	2	3
A state court	—	1
Each local classroom teacher	2	4
The National Institute of Education	1	3
The state legislature	—	15
A federal court	—	17
The U.S. Congress	—	25

*Twenty percent of all respondents incorrectly followed directions. Instead of listing only one agency trusted and one agency least trusted, they checked multiple agencies. These multiple response forms were not tallied.

- Local committees of teachers are most trusted in the South, and are twice as likely to be trusted by superintendents (25 percent) as by board members (13 percent).
- Southern respondents are decidedly less wary of state legislatures than are others and more than twice as wary of federal courts.

Another 5 percent wrote in "other" suggestions. Of these, four said "local control." One said "a committee of business personnel officers." The rest suggested combinations of agencies they would trust. Agencies listed in these combinations, in order of frequency, were: teachers, local boards, local citizens, parents, principals, administrators, superintendents, curriculum coordinators, students, business leaders, and university experts.

Nineteen suggestions were made about agencies that are least trusted. Of these, three mention HEW, two mention NEA, two say any federal agency, two say any legislature, two say any court, two say any state or federal agency, one said NAACP.

Why Competency Testing?

Asked why the competency testing idea appeals to people, respondents revealed consensus on two reasons: Most persons are attracted (1) by the aim to improve teaching and learning and (2) by the desire to make sure that every child is minimally competent. But beyond these reasons, school leaders recognize a diverse and contradictory set of motivations: to cut spending and to raise it, to prove schools good and to prove them bad, to cause curriculum change, to help minorities and to legitimize discrimination, and more.

About 60 persons offered other reasons, which can be categorized as follows:

Equity. To meet the needs of all students, give all students maximum opportunity, to make sure schools are doing what they are supposed to do: help all students.

Increase Learning. To help children, to see a child succeed, to make contributing citizens, to make students better adults, to improve the community education level, to help certain types to graduate (below average in "book intelligence"), to assure that young people can function in the world of today, to improve all skills to cope with the world of life or business.

Basics. To teach basics, to have confidence that schools are doing their job in the basics, to get back to basics, to change the thrust of mandated programs back to reading, writing, and arithmetic, to prove we must go back to the basics, to put more focus on academics.

Evaluation and Accountability. To provide accountability from the board, staff, students, and

facilities; to make sure children are getting the quality of education that the public is already paying for; to justify that tax dollars are being spent wisely and efficiently. To determine the competency of the school, to measure teacher effectiveness, to help evaluate our programs, to help determine the changing needs of instructional emphasis. To coordinate district goals, objectives, and programs; to help teachers focus their skills on specific objectives.

Standards. To make a diploma a more meaningful document, to assure that high school graduation is meaningful, to force standards and integrity back into education, to prevent the "passing on" of uneducated or poor students.

Motivation. To motivate students to at least a minimum competency—motivation by coercion if necessary.

Public Relations. To improve the credibility of education, to justify educational expenditures, to justify public education, to counter limited and narrow publicity citing extreme instances from distant situations.

Other. To discriminate against minority groups—actually to segregate. To provide a rationale for schools to have a last chance to attempt remedy. To prove their own opinions about education and to satisfy their own ego. Reaction to declining SAT scores. To join a fad.

Smaller, rural, and small town districts are more likely to believe that the reason is to justify more money for schools than are others. Otherwise, there are few notable patterns in the data but rather a remarkable consistency across categories of respondents.

Good and Bad Ideas

Asked their ideas about good or bad aspects of competency testing, school leaders were strong in their preference for "constructive" uses of such tests rather than using them to "get" incompetent students. A review of the items in Table VII shows that there is consensus nationwide about the value of giving a student several chances to pass a test, opportunity to prepare for a test, a series of tests throughout schooling rather than one test at the end of high school, and some measure of discretion in the use of test scores.

One would find disagreement nationwide about test secrecy, whether different competencies apply to different students, whether kids or schools should be able to use the tests voluntarily, and ways in which test scores might be published.

Unlike most other questions on the survey, several pronounced differences exist within categories of respondents.

- Superintendents and persons with more than three years' experience are more likely to believe that there should be different competencies required of persons of differing ability. Southerners are also more likely to believe this while Westerners are far less likely.
- Respondents from large, urban, and western districts are more likely than others to believe that students should know the kind of test items in advance of the test. They are also more likely to believe (along with superintendents and Easterners) that students should have several chances to take the test. Large urban districts and superintendents are also least likely to believe in test secrecy.
- Boards are more likely than superintendents to believe that school district participation in testing programs should be voluntary. And boards are more likely to favor publication of test scores than are superintendents, especially at the school district and individual school level.
- Midwesterners (at 61 percent) and Easterners (at 37 percent) differ over the importance of school district voluntary participation in competency testing. Easterners are also less likely (at 36 percent) than Midwesterners (at 49 percent) to want to keep test scores from being published.
- The larger the school district, the more likely it will want students to know the kind of test items in advance, to give several chances to repeat the test, and to want test scores published by district and by school.

Comments in this area from approximately 50 persons cluster around two points: half the comments acknowledge or warn that handicapped students must be treated specially. Other comments advise discretion in the uses of test scores and offer several suggestions:

- Only the general pattern of achievement should be publicized.
- Local districts should make information available to the media locally.

TABLE VII: GOOD AND BAD IDEAS

If minimal competency testing is established, which of the following characteristics do you think are good ideas? Which are bad ideas?

The typical respondent checked eight items which are good ideas and ten items which are not. Since almost every respondent checked every item, the "good" and "bad" responses add to 100%. Accordingly, only the "good" responses are reported here.

	Good ideas %
Tests are given *periodically* throughout K–12.	95
Tests are *required* of every student.	87
A student should know the *kinds* of test items in advance.	84
A student has *several chances* to repeat the test.	73
Tests should be secret.	54
Different competencies are required of students depending upon ability.	52
Tests are voluntary. *School districts* can decide whether to give it.	49
Test scores should be published for each school district.	47
The same competencies are required of all students.	46
Test scores should not be published.	43
Test scores should be published for each school.	34
A student has *one chance* to repeat the test.	26
Test scores should be published for each classroom.	17
A student should know the test items in advance.	14
Test scores should be published for each student.	9
Tests are voluntary. *Students* can decide whether to take it.	6
Tests are given near the end of high school *only*.	5
Tests are given only *once* to each student.	4

- If possible, each student should be given his or her score individually by a teacher or counselor.
- Test scores should be published on a state average basis, then each school can compare itself privately.
- Parents and students only should have knowledge of the test results.
- Test scores should be available to certain individuals who have a "need to know," but should not be published.
- Test scores should be available to the local board and administration, parents, teachers, and students, but not to the public.
- Secrecy is always bad.

Additional Comments

More than 60 persons accepted the invitation at the end of the survey to write additional comments.

Reaction to the concept of minimal competency testing ranged from unabashed enthusiasm to utter opposition. A suburban Illinois school board president wrote that "every school system owes to its community the obligation of saying 'This we guarantee to do for your child.' We owe to the child a good training in the basic skills, i.e., minimal competency." The opposite view was presented by a board president from suburban Pennsylvania: "Minimal competency testing is as worthwhile as a third horn on a bull or a second tail on a peacock . . . worthless and redundant."

The vast majority of comments, however, fell somewhere between these extremes. By far the most frequent attitude was one of approval with reservation, summed up by a board president from a suburban New York State district: "The idea of competency testing is sound. The major purpose should be to discover *early* (6th-9th [grade]) in what areas a student is deficient. Remediation in those areas should begin at earliest time." Or this comment from a suburban California board president: "I'm generally in favor of this, provided alternatives are provided to achieve these competencies and alternative testing methods are also available. . . . Minimum competencies in the 'basics' are just a beginning. If schools get bogged down there we may produce graduates who lack many competencies to function in society."

A high percentage of respondents expressed the opinion that the tests should be best seen as a

"tool" for setting high standards, diagnosing problems, and consequently improving teaching, the curriculum, and the education system. Many stated that tests should be the *means* not an end in themselves. As a diagnostic tool, then, most felt that the tests should *not* be used as a criterion for high school graduation, but rather as a way to uncover problems and provide remedial work for pupils who need it.

Despite the many reservations which prevented a wholehearted approval of minimal competency testing, there was a strong consensus of respondents that the high school diploma should mean something more than it now does. A board president from a small town in Vermont wrote that the high school diploma "should assure the reader that the holder has some low common denominator of knowledge and skill."

A recurring concern was the reservation about testing that "minimums" might become "maximums." A board member from Montana was skeptical about minimal competency testing as a requirement for high school graduation. "I cannot assume this," wrote that respondent, "and am disturbed by the tests I have seen which purport to measure such competencies. Testing for such low minimums does not justify in any way the billions of dollars spent every year in this country for the purpose of education." "Minimal competency standards imposed by either state or federal government will tend to establish a lowest common denominator which in a very short time will become the norm," wrote an Indiana board president from a large urban district.

Government interference at the state and federal level, and consequent erosion of local control, were cited frequently as a possible negative of minimal competency testing. "State or federally mandated minimum competency tests are another example of loss of local control or increasing mandates of the government," as a Minnesota board president put it. More strongly stated by an Iowa board president from an urban district, "Keep the federal government out of setting standards for competency—we already have TOO much federal control in the schools and are losing local control rapidly on oh, so many things in schools!" Another board president from Nebraska wrote, "Our district is working on minimal competency tests, really to beat a mandate by the legislature. We are going to try to avoid hurting students (and the school system) by them. We have an excellent system and would like to keep it that way." Other people felt that some involvement from the state was necessary. "I believe," wrote a small town Colorado board president, "that each state should decide if minimal competency testing is proper for that state. Then each local school district should define their own levels for minimal competency testing as they apply to that district." Two respondents supported federal involvement, including a board member from New York State, who said, "Standardized testing is a must if we are to *maintain* standards. State and national and local norms are important when evaluating curriculum and teacher-administration performance."

Many respondents felt that for their district the "problem" was already under control at the local district level. A board president from a suburban Ohio school district reported, "Over the past four years we have developed a set of district goals (arrived at by a committee of parents, students, teachers, and administrators and board) which are being translated into programs, goals, and objectives. It is at this point we have gotten into competency testing and minimum achievement levels. Some of the competency testing is in practice now, with more to be developed constantly."

Some people expressed the optimistic view that minimal competency testing will lead to better teaching and learning, but more were skeptical. "I find it difficult to see how a minimum competency test will force a change in the quality of education," said a suburban board member from Minnesota. "The result of testing and statistical data for defining weaknesses in a school district curriculum and teaching are available in most school districts. They just need to be used. Minimal competency testing will not force the issue, but merely cause additional expenditures that can be used better in other areas."

A few respondents expressed disdain for the additional "bureaucracy" of record keeping and the high cost of proposed minimal competency testing programs.

A small percentage of those who commented expressed reservations about minimal competency testing for another reason: that tests cannot pro-

vide an accurate portrayal of what a pupil really knows, that they might even distort the facts. One man feared that "competency testing will 'prove' a lot about the 'poor job' schools are doing, and I *do not* believe this is true." "Not all competencies are measurable," said another, "and not all persons should be placed up against the same standard." A person who reported having over 25 years of participation "in a vast scope of education as a graduate student, public school teacher, principal, and school board member" commented, "Students are not robots, automatons, nor computers to be programmed with a specific body of information. Tests are only one way to measure performance or mastery of skills."

Several people expressed the feeling that not enough attention is being given to the complexities of the issue. A board president from California felt that there was "too much rush toward implementation of such testing by legislators and state officials with little regard for the complexities of such a program; the diversity of the students to be tested, and the school districts where testing will occur; the consequences of denying a high school diploma to some students; and the costs of developing and implementing the tests." This person concluded by saying, "To ask if I favor having all students graduate with minimum skills is rhetorical. Of course. The question is, what will a particular testing proposal add to the effort."

In sum, board members and superintendents are sympathetic to the principles motivating the discussion of competency testing; they are not generally uncomfortable with testing as a school activity. But they are uneasy about several aspects of competency testing, probably have confused or divergent views on many issues, and certainly do not reflect a coherent set of opinions about this subject. Sensitive issues include the content of such tests, the purposes of their use, the pitfalls of their use politically, and questions of control. It seems clear that the ability to impose such tests gives strong control over curriculum. Board members tend to believe that local boards, community committees, or local staff are the best agencies to change testing rather than state or federal agencies, and especially rather than legislatures and courts.